12

CRM 89-129 / July 1990

Integration of Navy R&D Program Data: Are There "Families" of R&D Profiles?

Edward S. Cavin

CNA A Division of Hudson Institute
CENTER FOR NAVAL ANALYSES
4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268

Work conducted under contract N00014-87-C-0001.

REPORT DOCUMENTATION PAGE

Form Approved
OPM No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1990	Final

4. TITLE AND SUBTITLE
Integration of Navy R&D Program Data: Are There "Families" of R&D Profiles?

5. FUNDING NUMBERS
C - N00014-87-C-0001
PE - 65154N
PR - R0148

6. AUTHOR(S)
Edward S. Cavin

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Center for Naval Analyses
4401 Ford Avenue
Alexandria, Virginia 22302-0268

8. PERFORMING ORGANIZATION REPORT NUMBER
CRM 89-129

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Cleared for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This paper, undertaken as part of CNA's Quo Vadis II project, examines alternative statistical models for the cumulative distribution of cost and time of Navy R&D projects.

14. SUBJECT TERMS
Cost estimates, Costs, Distribution theory, Econometrics, Naval planning, Naval research, Optimization, Quo Vadis II, Regression analysis, Statistical models, Time

15. NUMBER OF PAGES
19

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
CPR	CPR	CPR	SAR

NSN 7540-01-280-5500

Standard Form 298, (Rev. 2-89)
Prescribed by ANSI Std. 239-18
299-01

CNA

CENTER FOR NAVAL ANALYSES

A Division of Hudson Institute   4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268 • (703) 824-2000

7 August 1990

MEMORANDUM FOR DISTRIBUTION LIST

Subj: CNA Research Memorandum 89-129

Encl: (1) CNA Research Memorandum 89-129, *Integration of Navy R&D Program Data: Are There "Families" of R&D Profiles?* by Edward S. Cavin, Jul 1990

1. Enclosure (1) is forwarded as a matter of possible interest.

2. This paper, undertaken as part of CNA's Quo Vadis II project, examines alternative statistical models for the cumulative distribution of cost and time of Navy R&D projects.

R. Kenneth Lobb

R. Kenneth Lobb
Director
Advanced Technology Program

Subj: Center for Naval Analyses Research Memorandum 89-129

Distribution List

SNDL		OPNAV
21A1	CINCLANTFLT	OP-00
21A2	CINCPACFLT	OP-00K
21A3	CINCUSNAVEUR	OP-02
24G1	COMSUBLANT	OP-03
24G2	COMSUBPAC	OP-05
28A1	COMCARGRU 4	OP-06
	Attn: RAdm R. MacKee	OP-07
A1	ASSTSECNAV RDA	OP-08
	Attn: Mr. R. Rumpf	OP-092
	Attn: Mr. J. Volpe	OP-094
A1	SECNAV	OP-094H
A1	UNSECNAV	OP-098
A2A	CNR	OP-22
A6	HQMC INTEL	OP-35
A6	HQMC PP&O	OP-50
A6	HQMC RD	OP-65
C4K	JPM-3	OP-70
E3A	NRL	OP-76
	Attn: Mr. P. Wilhelm	OP-940
	Attn: Mr. L. Hammerstrom	OP-943
FE1	COMNAVSECGRUCOM	OP-943E
FF44	NAVWARCOL	OP-987
FKA1B	COMSPAWARSYSCOM	
	Attn: RAdm. D. Brooks	
	Attn: RAdm. T. Betterton	
	Attn: Capt. F. Carey	
FKQ6C	NAVOCEANSYSCEN	
FKQ6F	NAVSWC DAHLGREN	
FKQ6G	NUSC	
FN1	COMNAVSPACECOM	
	Attn: RAdm. D. Frost	
	Attn: Dr. J. Springer	

CRM 89-129 / July 1990

Integration of Navy R&D Program Data: Are There "Families" of R&D Profiles?

Edward S. Cavin

Advanced Systems Division

CNA A Division of Hudson Institute

CENTER FOR NAVAL ANALYSES

4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268

ABSTRACT

This paper, undertaken as part of CNA's Quo Vadis II project, examines alternative statistical models for the cumulative distribution of cost and time of Navy R&D projects.

CONTENTS

ILLUSTRATIONS

TABLES

INTRODUCTION

In response to tasking from OP-98. CNA is undertaking a review of technology initiatives that will support the development of Navy systems[1] into the middle of the 21st century. The ultimate goal of this effort, Quo Vadis (Phase II). is to present OP-98 with an investment strategy that minimizes unevenness in Navy research and development expenditures over time and is consistent with projections of future R&D funding.

Developing such a strategy requires specification of a methodology for projecting future R&D project costs. Such a methodology should make best use of historical R&D project data to identify historical programs that are sufficiently similar to future projects to serve as a basis for projecting their cost profiles. Fortunately, most of the R&D programs being forecast by members of the Quo Vadis panel involve new generations of equipment functionally similar to systems already in being; one may reasonably suppose that the R&D cost profiles for these new systems will be similar to those of antecedent systems. For these future systems, the basic issue is *which* of the antecedent systems to use as a model, or perhaps how to construct an appropriate "average" cost profile from historical data.

Specifying models for cost profiles of substantially new systems is more challenging. In these cases, it will be necessary to identify commonalities in the technological requirements, system function, and platform support between the new system and historical programs. Ideally, one would like to develop a theory of R&D cost profiles, based on historical data, that could be generalized to substantially new situations.

This paper is an effort to specify models for R&D cost profiles of Navy systems. It examines the following as alternative methodologies for modelling cost profiles:

1) Average cumulative distributions of program cost and time[2];
2) Descriptive statistical models based on "curve fitting" to historical data; AND
3) Theoretical models of cost profiles estimated using historical data.

---

1. Throughout this paper, the term "Navy" will be understood to mean the Department of the Navy, i.e., the U. S. Navy and Marine Corps, including both active-duty and reserve forces.

2. Simple averages of program costs by time period are rejected out of hand for two reasons. (1) inflation effects are notoriously difficult to eliminate, and (2) such averages are too sensitive to be used to project scale and duration.

ALTERNATIVE APPROACHES

CUMULATIVE COST/TIME DISTRIBUTIONS

Cumulative cost/time distributions for historical R&D projects are the simplest models for new system cost profiles. Cumulative program costs are graphed as a funtion of cumulative project time elapsed, both in percentage terms.[1] Figures 1 through 3 show sample profiles for three different kinds of Navy systems: torpedoes, fighter aircraft, and air-to-air missiles (AAMs).

One can see from these figures that R&D cost profiles differ significantly among system types and often within syst.m types. Comparing figures 2 and 3, for example, it is clear that Navy fighter aircraft are developed on a much different schedule than AAMs. More than 75 percent of the total project cost for fighters is expended before 50 percent of the project time has elapsed, while for the average AAM, less than 50 percent of the cost is expended before half of the project time has elapsed. (Torpedoes tend to lie between these extremes.)

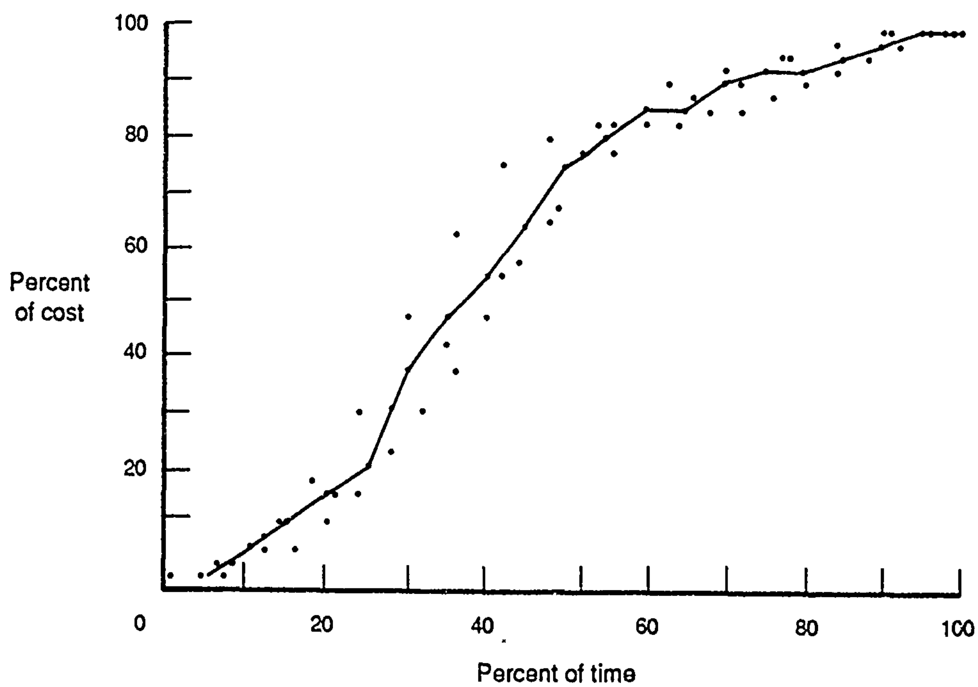

Figure 1. Cumulative cost/time curve for torpedo R&D

1. These profiles often are called "Lorentz curves" in the economics literature. They most commonly are used in studies of income distribution.

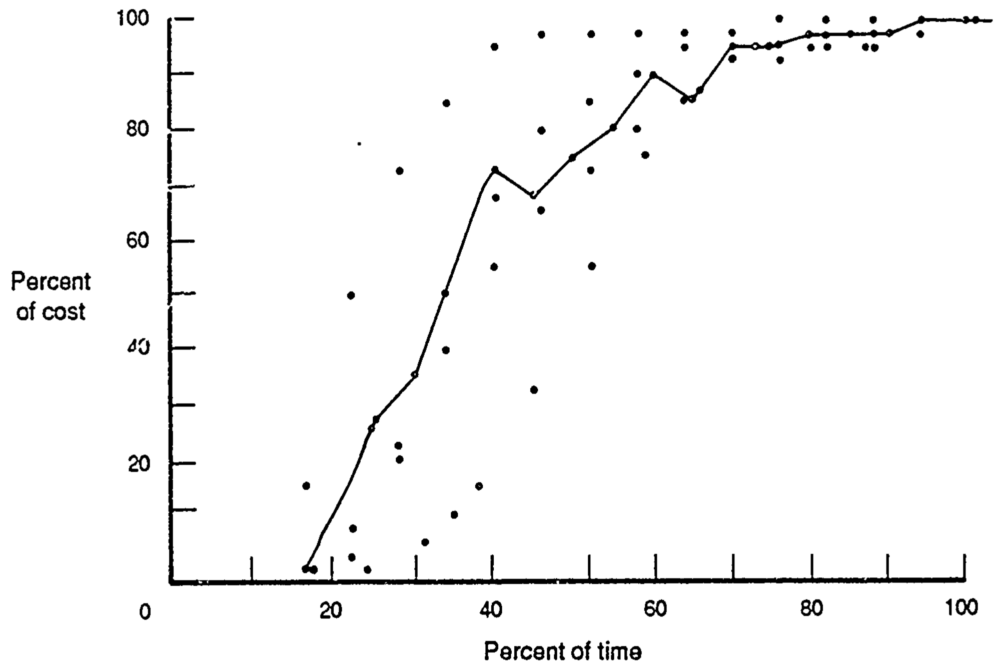

Figure 2. Cumulative cost/time curve for fighter aircraft R&D

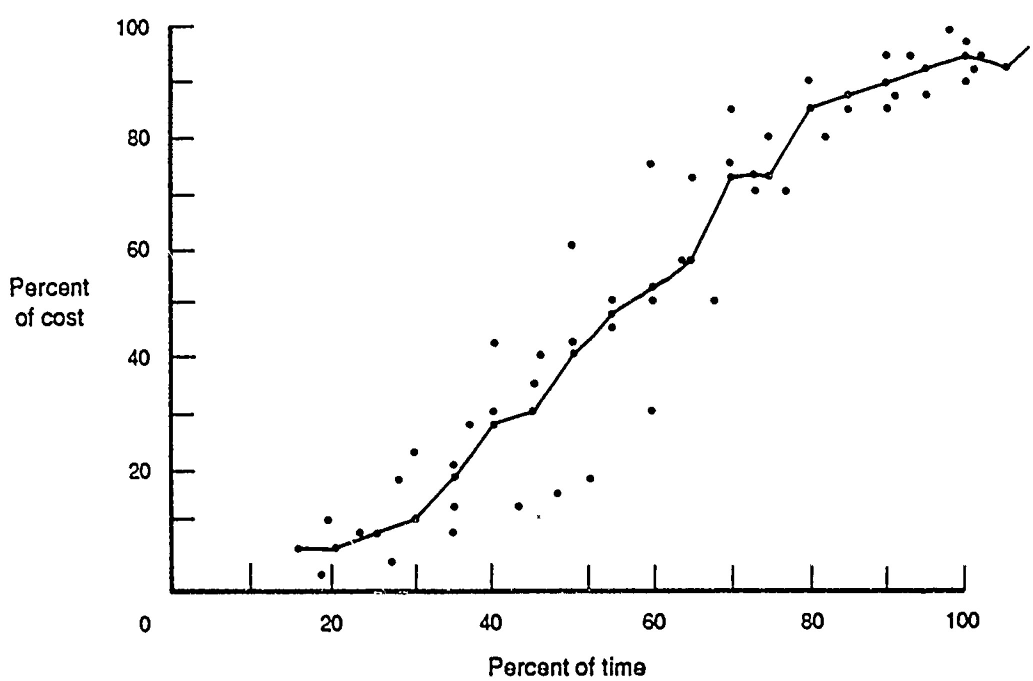

Figure 3. Cumulative cost/time curve for AAM R&D

It is tempting to use the mean cumulative cost/time distribution as the model for development costs of future systems.[1] However, figures 1-3 suggest that this must be done with caution. It appears from figure 1 that torpedo development is relatively generic, in the sense that several recent torpedo development projects have had almost identical cumulative cost/time distributions. Thus, the mean profile probably is a good choice for modeling follow-on torpedo developments. On the other hand, for fighter aircraft and AAMs, there is so much variation in cumulative profiles around the mean profile that it could be misleading to use the mean to model future developments. Because these development programs tend to be more idiosyncratic, it may be better to try to "match" the new system to be developed to a single antecedent system.

DESCRIPTIVE STATISTICAL MODELS

Another alternative for modeling R&D cost profiles is to estimate descriptive regression models of historical cost profiles. The concept of descriptive regression can be represented by the following equation:

$$\text{cost} = f(\text{time};\text{parameters}) + \epsilon \quad .$$

Alternative regression models are generated by specifying alternative functional forms for $f(.)$. Common choices are linear or polynomial models; however, since it generally is preferable to work with cumulative cost profiles, functional forms that represent probability distribution functions also should be considered. Two general distribution functions (in the sense that no symmetry is imposed) are the gamma distribution,

$$F(X) = \int_0^X \frac{\lambda}{\Gamma(\alpha)} y^{\alpha-1} \epsilon^{-\lambda y} dy \quad .$$

and the beta distribution,

$$F(X) = \int_0^X \frac{1}{B(p,q)} y^{p-1} (1-y)^{q-1} dy \quad .$$

For this analysis, the gamma and beta distributions, and a fourth-degree polynomial, were used as forms of the regression function. Since the gamma and

1. This is essentially how the NOSC "20/20" model constructs its "family curves."

beta distribution are nonlinear in the parameters, simple linear regression was not appropriate for these models. Instead, nonlinear least-squares methods were used.

Tables 1 through 3 present estimates for regression parameters of cumulative cost/time distributions for torpedoes, fighter aircraft, and AAMs. In each table, estimates are presented for polynomial, gamma, and beta specifications for the regression function. Each table also includes the standard R^2 goodness-of-fit measure for the alternative specifications.

Table 1. Estimated regression parameters for cumulative cost/time distribution (torpedoes)

Model/parameter	Estimated value	Standard error
Polynomial ($\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \alpha_4 x^4$):		
α_0	0.010	0.027
α_1	-0.753	0.403
α_2	10.253	1.717
α_3	-15.182	2.634
α_4	6.693	1.317
R^2=0.978		
Gamma distribution ($\int_0^X \frac{\lambda}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} dy$):		
λ	9.115	0.793
α	3.715	0.311
R^2=0.996		
Beta distribution ($\int_0^X \frac{1}{B(p,q)} y^{p-1}(1-y)^{q-1} dy$):		
p	3.884	0.068
q	6.116	0.068
R^2=0.989		

Table 2. Estimated regression parameters for cumulative cost/time distribution (fighter aircraft)

Model/parameter	Estimated value	Standard error
Polynomial ($\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \alpha_4 x^4$):		
α_0	-0.131	0.150
α_1	2.489	1.868
α_2	1.065	7.011
α_3	-5.670	10.019
α_4	3.270	4.810
R^2=0.815		
Gamma distribution ($\int_0^X \frac{\lambda}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} dy$):		
λ	8.724	2.171
α	2.558	0.602
R^2=0.966		
Beta distribution ($\int_0^X \frac{1}{B(p,q)} y^{p-1}(1-y)^{q-1} dy$):		
p	2.792	0.131
q	7.208	0.131
R^2=0.964		

Table 3. Estimated regression parameters for cumulative cost/time distribution (AAMs)

Model/parameter	Estimated value	Standard error
Polynomial ($\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \alpha_4 x^4$):		
α_0	0.070	0.101
α_1	-0.685	1.190
α_2	5.927	4.270
α_3	-6.281	5.873
α_4	1.953	2.719
R^2=0.934		
Gamma distribution ($\int_0^X \frac{\lambda}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} dy$):		
λ	7.526	1.039
α	3.753	0.496
R^2=0.980		
Beta distribution ($\int_0^X \frac{1}{B(p,q)} y^{p-1}(1-y)^{q-1} dy$):		
p	4.763	0.110
q	5.237	0.110
R^2 =0.967		

One can see that each of the regression models fits the cumulative cost/time distributions quite well. However, in each case the gamma and beta function specifications fit somewhat better than a simple polynomial, with the gamma function fitting slightly better than the beta. Figures 4 through 6 present the raw data and gamma distribution of fitted regression functions for the three systems.

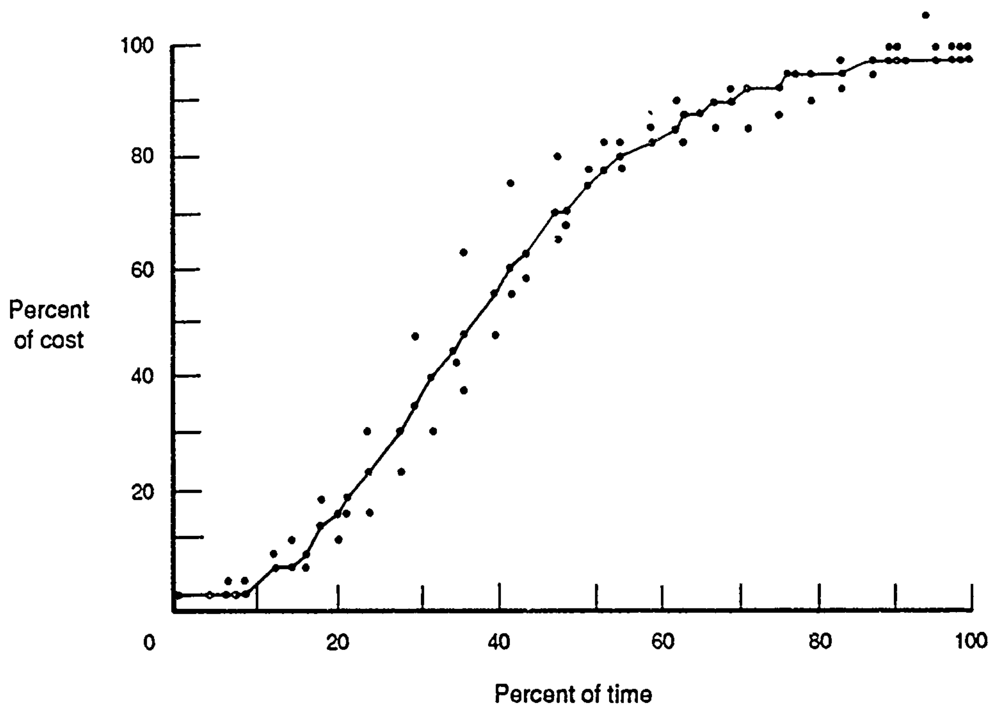

Figure 4. Descriptive regression data for torpedo R&D

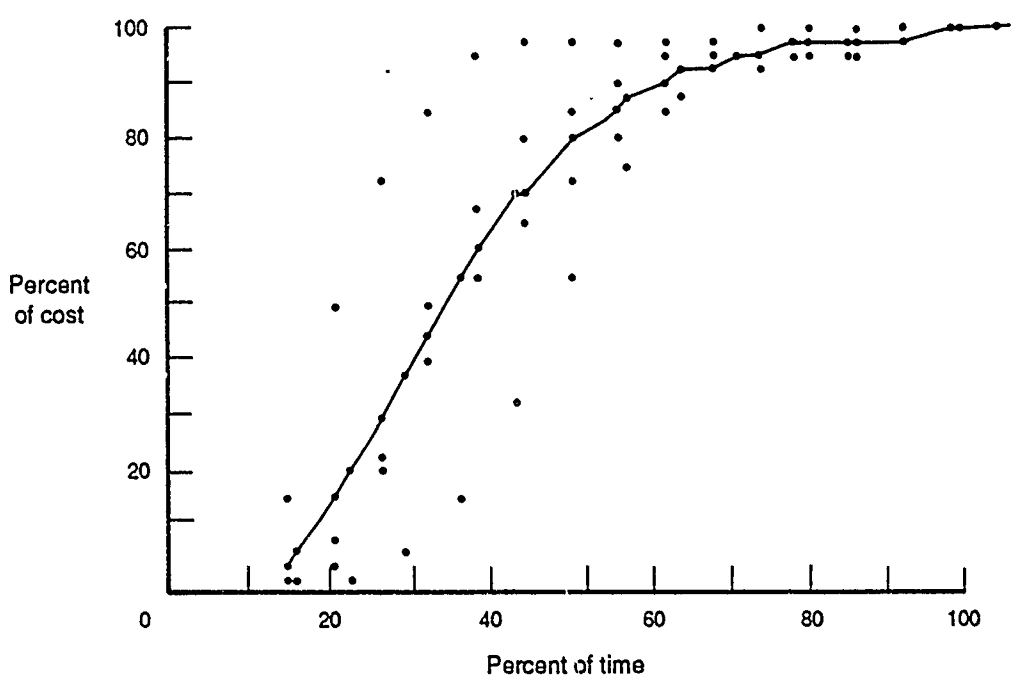

Figure 5. Descriptive regression data for fighter aircraft R&D

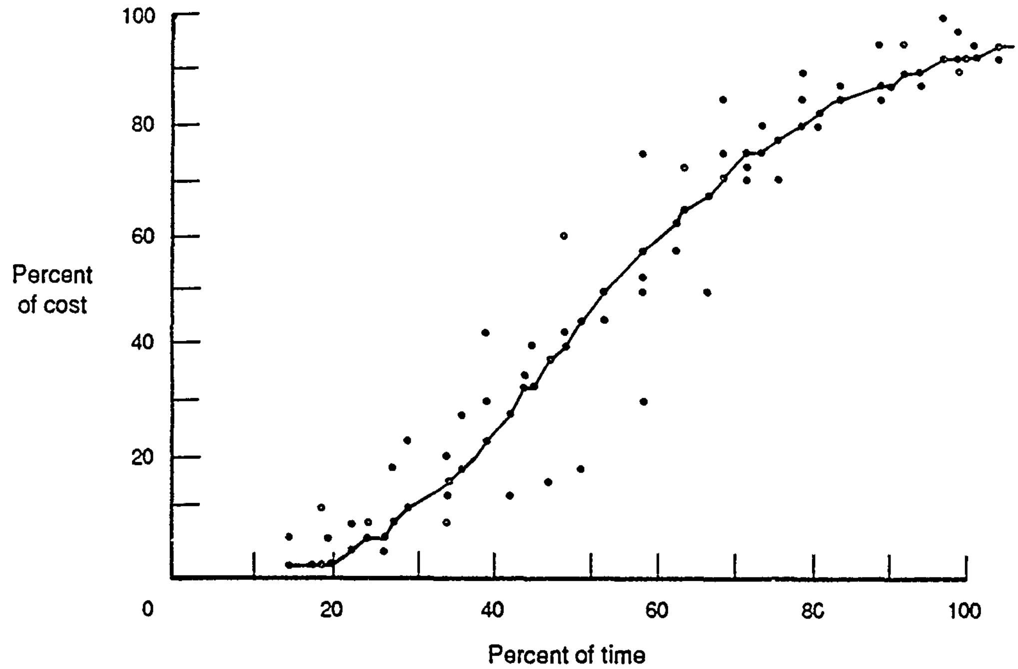

Figure 6. Descriptive regression data for AAM R&D

THEORETICAL MODELS FOR R&D COST PROFILES

Ideally, one would like to project future R&D cost profiles on the basis of a theoretical model of how R&D expenditures should be allocated for the development of a new system. As noted earlier, such a model is especially important if the new system is substantially unlike historical systems, so there is no obvious empirical model for the costs of the new system.

As a starting point for developing such a model, one needs to define the objective for spending R&D funds to develop a new system. Abstracting from the institutional facts of who actually makes programming decisions, suppose that the program manager needs to produce N units of a system by some time T hence. Presumably, he would try to do so at minimum total cost.[1] Thus, it seems natural to think of the problem in terms of classical optimization: minimizing the objective function (cost) within the constraints imposed by production requirements (N units) and available time (T years).

1. Issues of discounted costs will be considered subsequently.

However, the optimization problem as posed has a kind of intrinsic time dependence. The answer will be in terms of an optimal time path of R&D expenditures (and also of production costs). Therefore, it should be posed as a dynamic, as opposed to static, optimization problem. and it can be solved (if at all) using optimal control methods.[1]

Consider the following (simplified) optimal control problem:

$$\frac{max}{R(t), Y(t)} \int_0^T [C(Y(t), K(t)) - R(t)]dt$$

subject to $\frac{dX(t)}{dt} = Y(t)$, $\frac{dK(t)}{dt} = g(R(t))$, $X(0) = 0$, and $X(T) = N$. The variable $R(t)$ represents the time path of R&D expenditures, $K(t)$ the technical contribution to product purchased by $R(t)$, $X(t)$ the time path of system inventory, and $Y(t)$ the system production rate. The Hamiltonian for this control problem is:

$$H = [C(Y(t), K(t) + R(t)] + \lambda_1 Y(t) + \lambda_2 g(R(t)) \quad .$$

Suppose further that the functions $C(Y, K)$ and $g(R(t))$ have the following simple forms:

$$C(Y(t), K(t)) = \alpha ln(Y(t)) - \gamma ln(K(t))$$

$$g(R(t)) = BR(t)^{\beta}, 0 \leq \beta \leq 1 \quad .$$

The first assumption gives the cost function a familiar Cobb-Douglas form. The second assumption implies that the percentage rate of growth of $K(t)$ is less than linear in R&D expenditures, so that increasing R&D expenditures generate increasingly smaller increments to the rate of technology growth.

It can be shown that the first-order conditions for a maximum can be solved for the following optimal time paths:

$$\begin{aligned} R^*(t) &= At^{\frac{1}{1-\beta}} \\ Y^*(t) &= N/T \\ X^*(t) &= (N/T)t \quad . \end{aligned}$$

1. Good basic references on dynamic optimization are [1], [2], and [3].

Unfortunately, this solution for $R(t)$ is monotonically declining in t, which does not correspond very well to historical experience across systems. However, if the objective function is modified to represent the *discounted* system cost, the solution can be shown to follow the following optimal time path for $R(t)$:

$$R^*(t) = At^{\frac{1}{1-\beta}} exp(-\frac{r}{1-\beta}t) \quad ,$$

which has the correct generic shape. (The parameter r is the implicit discount rate.)

In fact, this function has the same general form as the gamma probability density function discussed above. This in turn suggests that the parameters of this simple theoretical model can be estimated using the gamma distribution function parameters estimated above, which turns out to be the case.[1] Tables 4 through 6 present estimates of the parameters of this model for torpedoes, fighter aircraft, and AAMs.

Table 4. Estimated regression parameters for structural cost/time model (torpedoes)

Model/parameter	Estimated value	Standard error
β	0.632	0.114
r	0.173	0.072
R^2=0.996		

Table 5. Estimated regression parameters for structural cost/time model (fighter aircraft)

Model/parameter	Estimated value	Standard error
β	0.358	0.386
r	0.367	0.346
R^2=0.966		

1. To avoid problems of program scale, it is convenient to transform the solution into a cumulative cost/time profile, which in this case will simply be the ratio of a gamma probability density function and its associated distribution function.

Table 6. Estimated regression parameters for structural cost/time model (AAMs)

Model/parameter	Estimated value	Standard error
β	0.637	0.180
r	0.228	0.091
R^2=0.934		

Estimates of the the discount rate and the rate at which R&D investments contribute to the accretion of project-specific technology are quite similar for torpedoes and AAMs. Unfortunately, the estimates of the parameters of the fighter aircraft model are too "noisy" to be much use. The estimated rate of technology growth in response to R&D expenditures is about 0.63, which implies that if R&D grows at 1 percent, the rate of technology growth will be about 0.63 percent. There is somewhat greater variability among the estimates of the discount rate parameter r, but neither of these rates (17 or 23 percent) seems intuitively unbelievable.

The lesson from this modeling exercise seems to be that it is easier to estimate the parameters of a model for more generic systems, such as torpedoes, than for less generic systems, such as fighter airplanes. If a modeling approach is to be used for aircraft R&D, it probably is best to use historical data for a similar kind of aircraft to estimate a model. Nevertheless, the simple control model outlined and estimated here suggests that important insights may be gained from a modeling approach.

COMPARISON OF ALTERNATIVE APPROACHES

It appears from the analysis undertaken in this paper that several alternative approaches may be acceptable for generating projections of R&D project cost/time profiles. The simplest method, averaging historical cost/time profiles, works best for systems that are highly generic, such as torpedoes. A slightly more sophisticated approach involves constructing a cost/time cumulative distribution that fits the observed data well in the least-squares regression sense. However, this approach also works best for systems that are highly generic. The final alternative investigated is to estimate the parameters of a theoretical cost minimization model using historical data, and use these parameter estimates to generate cost profiles for new system development. This third alternative is not demonstrably superior to the other two in terms of fit, but it may be attractive when projecting cost/time profiles for systems for which there is no close historical analogue.

As additional data become available, it will be interesting to estimate structural models and compare them to actual experience. If it turns out that there are stable empirical regularities in these projects, this information could help rationalize R&D planning, which, after all, is what the Quo Vadis effort is all about. In the meantime, the project staff should consider some kind of modeling approach as an alternative to the "first-round" practice of using the most recent (or conceptually most similar) system development project as a model, because using any single project as a model for all subsequent projects of a similar type is not an efficient use of historical information.

REFERENCES

[1] A. K. Dixit. *Optimization in Economic Theory.* Oxford University Press, 1976

[2] M. D. Intrilligator. *Mathematical Optimization and Economic Theory.* Prentice-Hall. 1971

[3] M. I. Kamien and N. L. Schwartz. *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management.* North-Holland. 1981